

Σ 6109:2

Fig. 1

GS wild-type

004020-29646450

10 20 30 40 50 60 70 80 90 100 110 120

130 140 150 160 170 180 190 200 210 220 230 240

250 260 270 280 290 300 310 320 330 340 350 360

370 380 390 400 410 420 430 440 450 460 470 480

490 500 510 520 530 540 550 560 570 580 590 600

610 620 630 640 650 660 670 680 690 700 710 720

730 740 750 760 770 780 790 800 810 820 830 840

850 860 870 880 890 900 910 920 930 940 950 960

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080

1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320

1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440

ATGAAATA ATATTATG ATATTATG ATATTATG ATATTATG ATATTATG ATATTATG ATATTATG ATATTATG ATATTATG ATATTATG ATATTATG

CTGCAAAAT GGTGTAATG TTGAAATAAC TTGTAATATA ATGATGCTGC TGCCTTTCCT TGCCTTTCCT TGCCTTTCCT TGCCTTTCCT TGCCTTTCCT TGCCTTTCCT TGCCTTTCCT

CCAACTGCTA CTGCTAAATTT AGTCACATAA TGTAAAGCTTA ATGCTGCTGC TGTAAAGCTTA ATGCTGCTGC TGTAAAGCTTA ATGCTGCTGC TGTAAAGCTTA ATGCTGCTGC TGTAAAGCTTA

AAATTTTATA ATGAAATGC TCCAAATTTT AATGCAAGTG CTAGTACATG CACAGCTTGT CACAGCTTGT CACAGCTTGT CACAGCTTGT CACAGCTTGT CACAGCTTGT CACAGCTTGT

TGTAAAGCTG CAGTCTCTAC TGTAAAGCTA CTGTAAGTG GAGTAACTAC TGAATATGTT AGATATCTCA CAGATATGTT TAAATATGTA CTTAACTTTT ACTATATATG TAAATATGTT

AATATCTCTT TCAATCCAGG TAAAGTTAA TSCACACCTT CTGCGGCAAT TAAAGCTT TAAAGCTT TAAAGCTT TAAAGCTT TAAAGCTT TAAAGCTT TAAAGCTT

TGCTCTGATG GTACTATAAG TGTCTCTGCA GTAAATATTT GGTAAAGCTA AATCACTGAA TGTAAATATTT GTGCTCTTAA CTTTAACTT AATATGCTC CTTAACTTCAA TCCAGCTTAT

AGTAACTGCC TACCTTGCCC AGCAATATAA GATTAATGCTG CTGAAAGCTC TCAAGCTGCT GCTGAAAGCT GCTGAAAGCT GCTGAAAGCT GCTGAAAGCT GCTGAAAGCT GCTGAAAGCT

GGAGCACTA ATTAATGTAAT ATTAATAACA GAATGCTCTA ATGCTGCTGC TAACTTTTAT TTGTAAGCTA ATAACTTCTA GCTGAAAGCT AGTAAAGCTA AGCTATGCTC AGCAATATAA

GTTTAAAGCG CTGTAAAGAC TCCAGCTGCT ACTGCTACTT TAAATGCTA ATGCTGCTGC GATGCTGCT GATGCTGCT GATGCTGCT GATGCTGCT GATGCTGCT GATGCTGCT

TCTAAATGCT TTAATGCTG TCCAACTTTT TAACTACAA AATAACTGA TTGCTGCTA GGTAAAGCTA GGTAAAGCTA GGTAAAGCTA GGTAAAGCTA GGTAAAGCTA GGTAAAGCTA

CCGAACTG CTAAATAAAA TAAATATGTT GATTTGCTA ATTTTATCT AATTTCTT TAAATATGTT TAAATATGTT TAAATATGTT TAAATATGTT TAAATATGTT TAAATATGTT

Coding region: nucleotides 1-1404

Fig. 2(a)

## G5 SYNTHETIC

004020" 29626460

SE410  
NU:  
102

10	20	30	40	50	60	70	80	90	100	110	120
ATGAGAACA	ACATCTGCT	GATCTGATC	ATCTCTCTG	TCATCAACA	GATCAGCTT	GCTAACCTG	CTGTGCGAC	ACCGCTGGC	AGCTGACGA	CCTGGGAAC	
130	140	150	160	170	180	190	200	210	220	230	240
CCTGCTAAC	GTGTGATCT	TCAGAGAAC	TTCTACTAC	ACAAGCTGC	TGCTTTCTG	CCTGGAGCT	CTTACCTGAC	CCCTTTGCT	CAGAGAGAG	ACGCTGZAG	TCAGCTTAC
250	260	270	280	290	300	310	320	330	340	350	360
CCCTCTGCT	CCGCTAACCT	GCTGACCCG	TGTAACTGA	AGTGTCTGC	TGGAACGCT	ATGCTGCGG	GAGCTACCG	CTACGCTCT	ATCATCACCG	AGTGTGTGA	CTGTGCTAT
370	380	390	400	410	420	430	440	450	460	470	480
AACTTCTAC	ACGAGAGGC	TCTTAATCT	AACGCTGGG	CTTCTACCT	TACCGCTGT	CTGTGACCC	CGCTGGAGG	AGCTGTGAC	GCTGAAAGG	CTGCTACCAT	CGTGGCTGAG
490	500	510	520	530	540	550	560	570	580	590	600
TGTAACTGG	CTTGTCTAC	CGGAACGCT	CTGGACGAG	GAGTACAC	CGACTACCT	CGCTTTTCA	CGGAGTGTG	GAGGTGTGC	CTGAACTCT	ACTTAAACG	AAACAAACGA
610	620	630	640	650	660	670	680	690	700	710	720
AAACACCTT	TCAAACCTG	AAAGTCTCAG	TGTAACTCT	GTCTCTCTAT	CAAGCTGCT	AAAGTGGCT	AGGCTACCT	GGGAAACGAC	GCTAACATCA	CCGCTGAGT	TACGCTGCT
730	740	750	760	770	780	790	800	810	820	830	840
TCCTCTGAG	GAACTCTC	TCCTCTGGA	GTAACAAT	GGTGGCTCA	GGACACGAG	TGTAACTCT	GTCTCTTAA	CTTCTACAC	AACAACTCT	CTTAACTTCA	CCCTGGAAC
850	860	870	880	890	900	910	920	930	940	950	960
TCTACTGTC	TGCTCTGTC	TGCTAACAG	GACTACGAG	CTGAGCTAC	CGCTGAGGA	GCTGCTAC	TGCTTACCA	GTGTAACTC	GCTGTGCTG	ACGGAACGC	TATGCTTCT
970	980	990	1000	1010	1020	1030	1040	1050	1060	1070	1080
GGAGCTACA	ACTAGTCT	CCTGAGACC	GAGTGTCTA	ACTGCTGCT	TACTTCTAC	TTGAGCGGA	ACAATCTCA	GCTGCTGAT	TCTGCTGTA	AGCTTGTCT	TGCTAACAG
1090	1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200
GTGAGGAG	CTGTGCTAC	CGCTGAGGA	ACCGCTAC	TGATGCTCA	GTGTGCTCT	GTGTGCTCT	CTGAGCTCT	GCTAACGAG	GGAACTACT	CTTAACTTCA	GTGAGCTCT
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320
TCGTAGTCT	TGAAGTCT	TGCTACTTC	TACACCTCA	AGCAGACGA	CTGCTGCT	GGAATGACA	CGTGTACCT	TGTAAACAG	AACTGACT	CTGAGCTCA	GGCTAACCT
1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440
CGTGAATCT	CTAAGAGGA	CACTGCTGT	GACTTCTCT	ACTTCTGCT	TACTTCTCT	CTGCTGCT	CTTAACTCT	CTTAACTCT	CTTAACTCT	CTTAACTCT	CTTAACTCT

Coding region: nucleotides 1-1404

Fig. 2(b)

Sequence alignment of 48 kD G1 i-antigen  
and 55 kD G5 i-antigen protein sequences

Seq ID No: Conserved regions

Seq ID No: 6	G1	1	MKYNILLIISLFINELRAVPCPDGTQ- <u>QAG-LTDVGAADLGTCVNC</u> RPNFYNN----	90	MKYNILLIISLFIN
Seq ID No: 7	G5	1	MKNILVILLISLFINQIKSANCPCVGTETNTAGQVDDLGRP-- <u>ANCVCQKQNFYNNAAA</u>	91	MKNILVILLISLFIN
			** ***: *****		
	G1	55	--GGAA-----QGEANGNOP-		
	G5	54	FVPGASTCTPCPOKKDAGAPNPAPANLVTQCNCVKCPAGTALAGGATDYAAIITECVNC		
			**:		
	G1	68	---P-----AAN-NRARGICVPCQINRVGSVTNAGDLATLATQCSTQCTGTALDDGVTDV	92	CPTG-TALDDGVT
	G5	119	RINFYNENAPNPNAGASTCTACPNVRVGGALTAGNAATIVAQCNVACPTGTALDDGVTDV		
			* ** **		
	G1	120	FDRSAAQCVKCKPNEFYNGSGSPQGEAPGVQVFAAGAAAAGVAAVTSQCVPQCQLNK--NDS	93	CVKCKPNFYNYNG
	G5	179	YVRSFTQCVKCRNFYNGNN--GNTP-----FHPGK-----SQCTPCPAIKPANVA	94	CVKCLNFYNYNG
			: ** **		
	G1	178	PATAGAQAANLATQCSNQCPTGTVLDDGVTLVNTSATLVCVKRPNFYNGSGSPQGEAPGV		
	G5	214	QATLGNDATITACQNVACPDGTISAAGVN-NWVAQNTCTNCAPNFYNN-----AP--		
			** * : **		
	G1	238	QVFAAGAAAAGVAAVTSQCVPQCQINKND-SPATAGAQAANLATQCSNQCPTGTAIQDGVTL	95	CPAGTVLDDGT
	G5	295	-NPNPG-----NSTCLPCPANKDYGAEATAGGAATLAKQCNACPDGTAIASGATN	96	CPAGTVLTDGT
			* *		
	G1	317	VFSNSSTQCSQCIANYFENG-NFEAGKSQCLKCPVSKTTPAHAP-GNATQATQCLTQCP		
	G5	325	-YVILQTECLNCAANFYFDGNWFQAGSSRCACAPANKVQGAATAGGTATLIAQCALECP		
			: ** * : **		
	G1	355	AGTVLDDGTSTNFVASATECTKCSAGFFASKTTGFTAGTDCTCTECTKLTSGATKYVAE	97	AGTDCTECTKKLTS6AA
	G5	384	AGTVLTDGTSTYKQAASECVKCAANFYTTKQTDWVAGIDTCTSCNCKLTSGAENLPES	98	AGTDCTCTSCNKKLTSGAEA
			**** **		
	G1	415	ATQKVQCASTTFAKFLSILLFISFYLL	99	FAKFLSISLLFISFYLL
	G5	444	AKKNIQCD--PANFLSILLISFYLL	100	FANFLSISLLISFYLL
			* : **		

Fig. 3(a)

THE UNIVERSITY OF CHICAGO

G1 G5 GTTCCAGTCTCTGATGGTACTAGACTCA-----AGCTGGAT-----TGACTGATGTAGGTGC  
 GCTAATTOTCTCTGTGGAACTGAAACTAACACAGCCCGATAAATGTA-TGATCTAGGAAC  
 .....  
 G1 G5 TGCTGATCTTCGACTTTOTGTTAATTGC-AGACCTAATTTTTACTATAATGTTGTGCTG  
 TCCT-----GCAAATGTGTGTTAATTGTTAGAAA-AACTTTTATTATAAATAATGCTGCTG  
 \* \*\* \* \* \*\*\*\*\*  
 G1 G5 CTTAAGGAGAAGCTAATGGTAATTAACCTTTGCGCAGCAAAATAATGCTGCTAGAGGTAAAT  
 CTT-----TCGTTCC-----TGGTGCTAG---TACGT  
 ... ..  
 G1 G5 GTGTACCATTG-CCA-AATAAACAGA-GTAGGCTCTGTTACCAA-TGCAGGTG--ACTTAG  
 GTACACCTTTGTCCATAAAAAAAAGATGCTGTGCT-TAACCAATCCACCTGTCTACT--G  
 \*\* .....  
 G1 G5 CTACTTTAGCCACATAATGCGAGTACTTAATOTCTCTACTGGCACTGCACTTGATGATGGAG  
 CTAATTTAGTACATAAATGAAGCTTTAAATGCCCTGCTGGTAGCCCAATTGCCAAGTGAAG  
 \*\*\* \*\*\*\*\*  
 G1 G5 TGACAGATGTTTTTG--ATAGATCAGCCGCATAATGCTTTAAATGCAAACCTAACTTTTA  
 CAACAGATTAAGCAGCAATA-ATCA-----CAGAATGTGTTAATTGAGAAATTAATTTTTA  
 \*\*\*\*\*  
 G1 G5 CTATAATGOTGGTTCTCCTTAAGGTGAAGCTCCTGGCGTTTAAGTTTGTGCTGCTGGTGC  
 ---TAATGA-----AA-----ATGCTCC-----AAATTTTAA-----  
 \*\*\*\* ..  
 G1 G5 TGCCGGCTGCAGGTGTGCTGCGGTACTACTTAATGTTGACCTTGCCAACTAAACAAAAA  
 -----TGCAGGTG-----CTAGTACATGCACAGCTTTGTCCGGTAAACAGAGT  
 \*\*\*\*\*  
 G1 G5 CGATTCTCCTGCCACTGCGAGT---GCCTAACTAATTTAGCCACATAATGTAGCAATTA  
 TGTGTTGTCATTGACTGCTGTTAATGCC---GCTACCAATGTCGATAATGTAAAGCTGCC  
 .....

Fig. 3(b)

G1 ATGTCTACTGGCACTGTACTTGATGATGGAGTGACACTGTGTTTAAATACATCAGCCAC  
G5 ATGTCTACTGGTACTGCACTTGATGATGGAGTAACTACTGATTATGTTAGATCAATTCAC  
.....

G1 ATTATGTGTTAAATGACAGACCTAACTTTTACTATAATGGT-----GGTT---CTCCCTA  
G5 AGAATGTGTTAAATGTAGACTTAACTTTTACTATAATGTTAAATGGTAATACTCCCTT  
.....

G1 -----AGGTGAA-----GCTCCTGGGGTTTA  
G5 CAATCCAGGTAAGTAAATGCACACCTTGTCCGCAATTAAACCTGCTAATGTTCTTA  
.....

G1 AG-----TTT-----TTGC-----TGCTGG  
G5 AGCTACTTATAGGTAATGATGCTACAATAACCGCATATGTAACGTTCATGCCCTGATGG  
.. ..

G1 TGCT-----GCCGCTGCAG-----GTGTTGC-----  
G5 TACTATAAGTCTGCTGAGTAAATTAATGGGTAGCACAAACACTGAATGTACTAATTG  
.. ..

G1 -----TGCCCTTACTAGTTAATGRT  
G5 TGCTCCTAACTTTTACAATAATAATGCTCCTAATTTCAATCCAGGTAATAGTACATGCC  
.. ..

G1 ACCCTGCCAATAAACAACAAACGATTCTCTG---CCACTGCAGGTGCCTAAGCTAATTT  
G5 ACCTTGCCCAAGCAATAAAGATTATGGTGTGAAGCCACTGCAGGTGGTGGCGCTACTTT  
\*\*\*\*\*

G1 AGCCACATAATGCAGTACTTAAATGTCCAACCTGGCACTGCAATT-CAAGACGGAGTGACAC  
G5 AGCCAAATAATGTAATATGTCATGCCCTGATGGTACTGCAATTGCTAGT-GGAGCAAC--  
\*\*\*\*\*

G1 TTGTTTTAGTAAT-TCATCCACATAATGTTCTTAAT-GCATTGCTAATTACTTTTTAA  
G5 -TAATTAT-GTAATATTATAACAGAAATGT-CTAAATTGTGCTGCTAACTTTATTTTGA  
.. ..

G1 TGGTAAT---TTCAAGCAGGTAAAGTTAATGTTTAAAG--TGTCCAGTAAGTAAACT  
G5 TGGTAATAATTTCTAGGCAGGAAGTAGTAGATCC--AAAGCATCTCCAGCAATAAAGTT  
\*\*\*\*\*

G1 A-----CTCCAGCACATGCTCCAGGTAATGCTACTTAAAGCCACATAATGT----TT  
G5 TAAGGCGCTGAGCAA---CTGCAGGTGGTACTGCTACTTTAATTGCATAATGTGCCCTT  
.. ..

G1 GACCACATGTCCTGCTGGTACAGTACTTGATGATGGAACATCAACTAATTTGTAGCTTC  
G5 GA-----ATGCCCTGCTGGTACTGTACTCACCAGTGAACAACATCTACTTATAAATAAGC  
.. ..

G1 CGCAACTGAATGTACTAAATGTTCTGCTGGCTTTTTCATCAAAAAACAACCTGCTTTAC  
G5 AGCATCTGAATGTGTTAAATGTGCTGCCAAGTTTATCTACAAAAATAAAGTATTGGGT  
.....

G1 ACCAGGTACTGATACATGTACTGAATGTACTAAAAAATAACTTCTGGTGCCACAGCTAA  
G5 AGCAGGTATTGATACATGACTAGTTGTAATAAAAAATAACTTCTGGCGCTGAAGCTAA  
\*\*\*\*\*

G1 AGTATATGCTCAAGCTACTCAAAAAG---TATAATGCGCCTCCACTACTTTTCGCTAAAT  
G5 TTTAC---CTGAATCTGCTAAAAAATAATATAATGTG-----ATTTCCTAATTT  
.. ..

G1 TTTATCGATTTCCTTATTATTATTCTTTCTATTATTG  
G5 TTTATCAATTCCTTATTATTGATTCTTATTATTATTA  
\*\*\*\*\*

Fig 3(b)  
(cont)

# 65 kD i-antigen protein

10 20 30 40 50 60  
 E0 1040:7 MKNNILVILI ISLFINQIKS ANCPVGTETN TAGQVDDLGT PANCVNCQKN FYYNNAAFV  
 70 80 90 100 110 120  
 PGASTCTPCP QKRDAGAQP N PPATANLVTO CNVKCPAGTA IAGGATDYAA IITECVNCRI  
 130 140 150 160 170 180  
 NFYNENAPNF NAGASTCTAC FVNRVGGALT AGNAATIVAQ CNVACPTGTA LDDGVTTDYV  
 190 200 210 220 230 240  
 RSFTECVKCR LNFYYNGNNG NTPFNPFGKSQ CTPCPAIPKA NVAQATLGND ATITAQCENVA  
 250 260 270 280 290 300  
 CPDGTISAAG VNNWVAQNT E CTNCAPNFYN NNAPNFNPGN STCLPCPANK DYGAETAGG  
 310 320 330 340 350 360  
 AATLAKQCNI ACPDGTALAS GATNYVILQT ECLNCAANFY FDGNNFQAGS SRCKACPANK  
 370 380 390 400 410 420  
 VQGAVATAGG TATLIAQCAL ECPAGTVLTD GTTSTYKQAA SECVKCAANF YTTKQTDWVA  
 430 440 450 460 470 480  
 GIDTCTSCNK KLTSGAEANL PESAKKNIQC DFANFLSISL LLISYLL\*\* .....

004020 29626460

Fig. 4

48 kDa G1 i-antigen repeats

SEQ ID NO:

	10	20	30	40	50	60	70	80	90	
8	CPDSTOTQNS	ITDVLAAADLG	TCVSLRPNEY	YEGGAOGEA	NGNDFAPEN	AAKCI	CVPCQSRV	GVVIRAGLA	ELATQCST	Q
9	CPTGTALDDG	VITVIDESAA	QCVKCRPEFY	YNGGSEGEA	FGVVFANGA	AAAGVAAVTS	QCVPCLERN	DSPATAGACA	ELATQCSN	Q
10	CPTGTALDDG	VITVFTSAT	LCVLCRPNEY	YNGGSEGEA	FGVVFANGA	AAAGVAAVTS	QCVPCLERN	DSPATAGACA	ELATQCST	Q
11	CPTGTALDDG	VITVFTSAT	QCSQCLAYF	ANGF	FEA	GR	SQCICFVST	TPAAGGT	ATCATOCT	T
12	CPAETVIDDG	TSTNEVMSAT	EDKACSAGFF	ASKTTG	FPA	GTD	TCTICTKLT	SLATKVMA	EATOKVOCAS	T

\* \* \* \*

Fig. 5(a)

004020-29626460



# 55 kDa G5 i-antigen repeats

Seq 10 no.	55	1	CPVGTETNTAGQVDDLGTPANCVNCQKVFYVNNAA---AFVPGFSTCTPCPKKDAEAGPNPPATNLT-OCNVK
	56	2	CPAGTALGG-AIDYAAIITECVNCRINFYNAP-----NFNAGASTCTACHVNRVGGALTAGNAATIV--AOCNVA
	57	3	CPTGTALDDGVITDYVRSFITECVRQRINFVYVNGNNGNTBFNPGKSCCTPCPAIKPANVAQATLGNDAITITAQCENVA
	58	4	CPDGTLSAAGVNNWVAQN-TECTNCAFPNFYANNAP-----NFNPGNSTQPCPANKDYGAETACGCAATLAK-OCNFA
	59	5	CPDGTALASGAFTNYVILQ-TECLNCAANFYFDGN-----NFQAGSSRCKACPANKVQEAAPAGGIVATLI-AQCALE
	60	6	CPAGTVLTDTGTHSTYKQAASECVKCAANFYITTKQ---TDWVAGIDICTSCNKKLTSGAANLPESAKKNI-QCDFA

Fig. 5(b)

004020-29626460

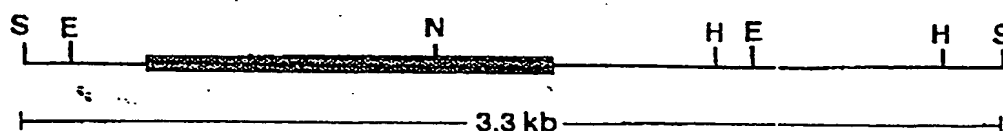


Fig. 6

Diagram illustrating the G1 isolate genome organization. The cDNA sequence is 1185 nt long, starting with an ATG start codon and ending with a TGA stop codon. The GENE sequence is 1925 nt long, also starting with an ATG start codon and ending with a TGA stop codon. A large arrow points from the GENE to the text "48 kDa antigen (G1 isolate)".

Diagram illustrating the alignment of cDNA and gene sequences. The top line represents the cDNA sequence, with positions 897 and 1172 indicated. The bottom line represents the gene sequence, with positions 56, 952, and 1227 indicated. A vertical bracket connects the two lines at position 952.

63 V N I H O .  
 1173 GTAAATTCATTAATCAAGCTTCGAAAAACATGCTGCTAGTACCTTATCTAGCTTGAAGTATTTAATCAAGA;  
 65 K V Y A E A T G K V C C A S T T F A K P L S I B L I P I S F Y L L .  
 1228 AAGTATATGCTGAAGCTTACTCAAAAAGTAAATGCGGCTCCACTACTTTGCGTAAATTTTATGATTTCTCTATATTTATTTCTTCTATTTATTTGTA  
 67 AAGTATACGCTGAAGCTTACTCAAAAAGTAAATGCGGCTCCACTACTTTGCGTAAATTTTATGATTTCTCTATATTTATTTCTTCTATTTATTTGTA  
 68 AAGTATACGCTGAAGCTTACTCAAAAAGTAAATGCGGCTCCACTACTTTGCGTAAATTTTATGATTTCTCTATATTTATTTCTTCTATTTATTTGTA  
 69 AAGTATACGCTGAAGCTTACTCAAAAAGTAAATGCGGCTCCACTACTTTGCGTAAATTTTATGATTTCTCTATATTTATTTCTTCTATTTATTTGTA  
 +1409 +1413  
 1330 TGAATAAATTAATTCATATCTACTTATTTTATTTATTTATGTTTATAAATGAAAAATAGATAAAATTTAAATATTTTAAATTAATTTTATATTA  
 TGAATAAATTAATTCATATCTACTTATTTTATTTATTTATGTTTATAAATGAAAAATAGATAAAATTTAAATATTTTAAATTAATTTTATATTA  
 TGAATAAATTAATTCATATCTACTTATTTTATTTATTTATGTTTATAAATGAAAAATAGATAAAATTTAAATATTTTAAATTAATTTTATATTA  
 TGAATAAATTAATTCAT

Fig. 7

20 AVFPDGTQTOAG-LTVDVGADIGTCVNC-RPNFYNGNAQG---ZANGNQP-FAANNAAR---GICUP---CQINRVGSVTVNAGDLAT  
5'440 W01.61 i-ag  
61 AVDCQ-----GSAGYTTDDSEVDAKCKCNAPCTACAGTADKCTKCDANGAAPLYLKTNPSPDTGTCVSAVDCQG-SAGYTT--DDSVSD  
5'640 W01.62 vspA6-S1  
20 AVFPDGTQTOAG-LTVDVGADIGTCVNC-RPNFYNGNAQG---ZANGNQP-FAANNAAR---GICUP---CQINRVGSVTVNAGDLAT

i - a g  
v s p a 6 - 5 1

i-a9  
vspA6-S1

i-ag  
vpa6-S1

i-ag  
vspA6-S1

File 00

004020" 29626460

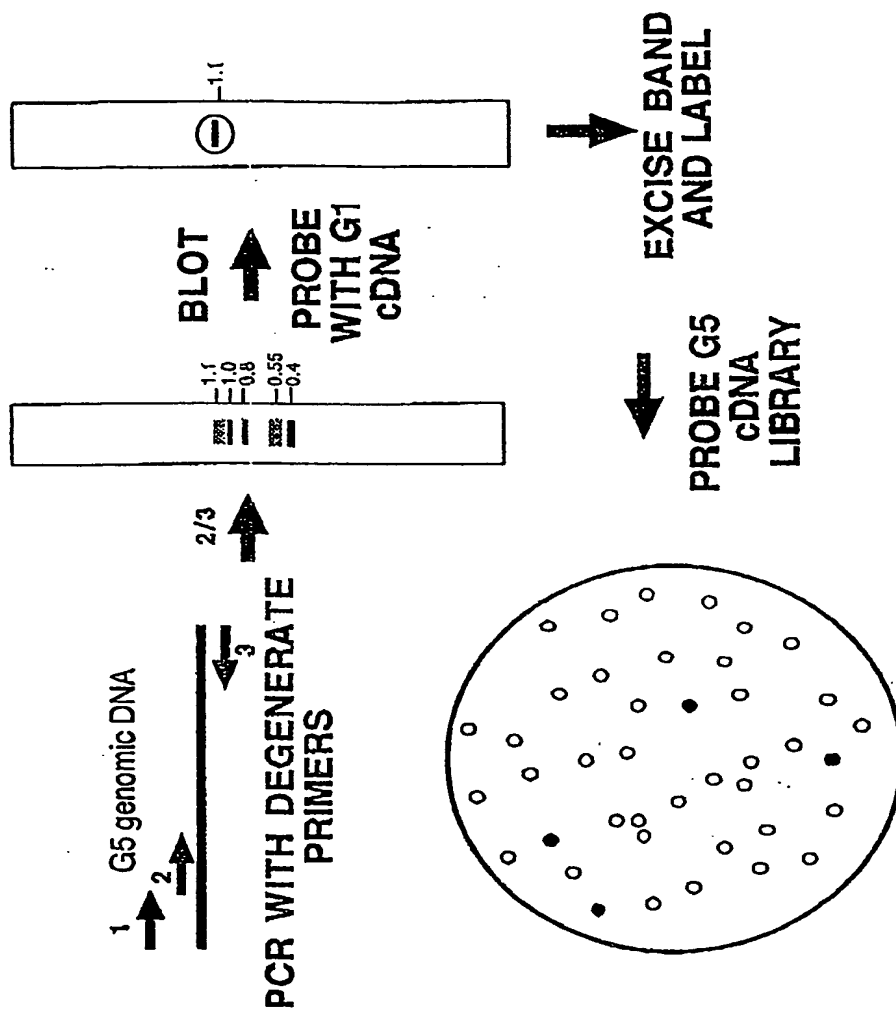


Fig. 9

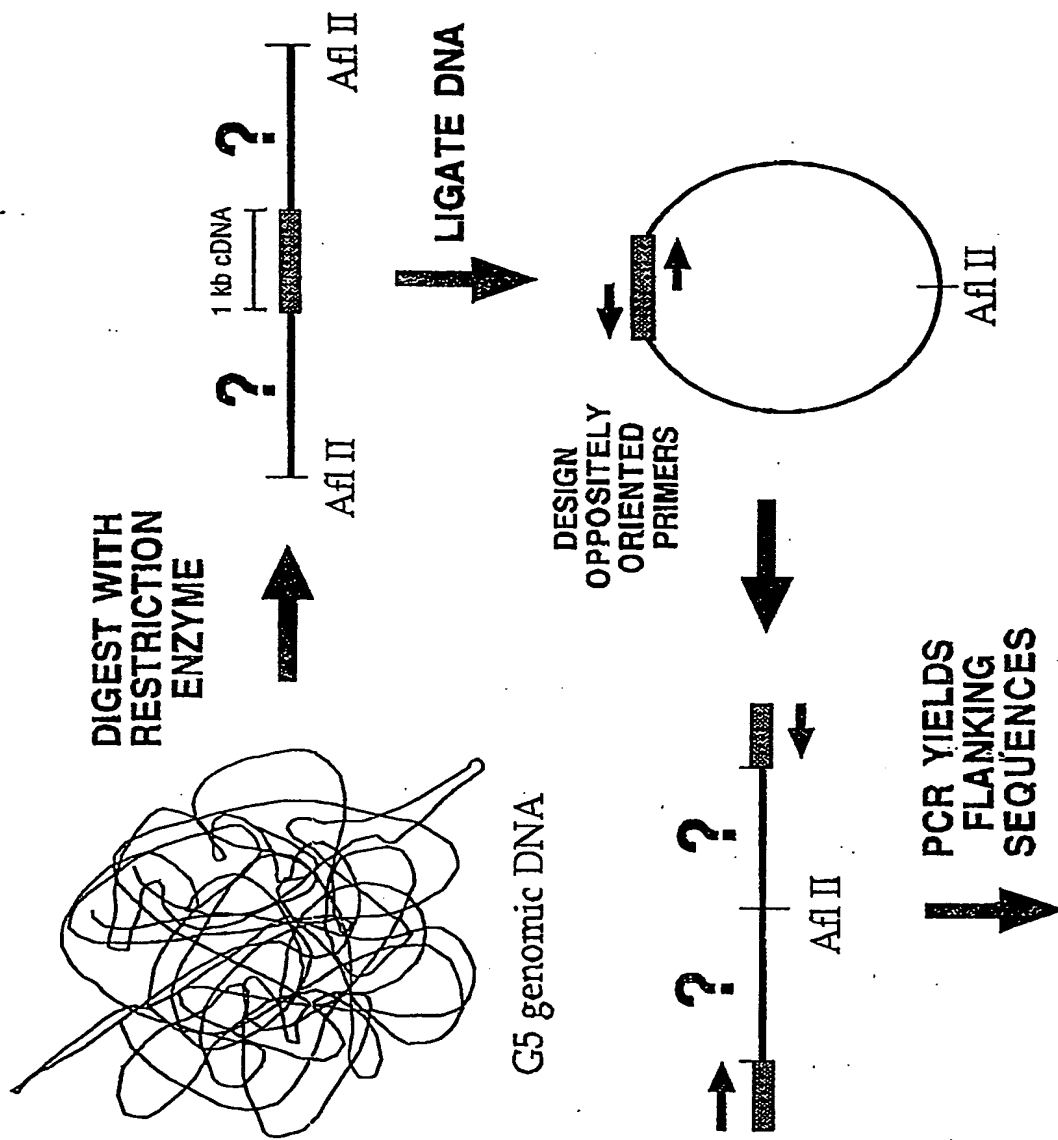


Fig. 10

004020-29626460

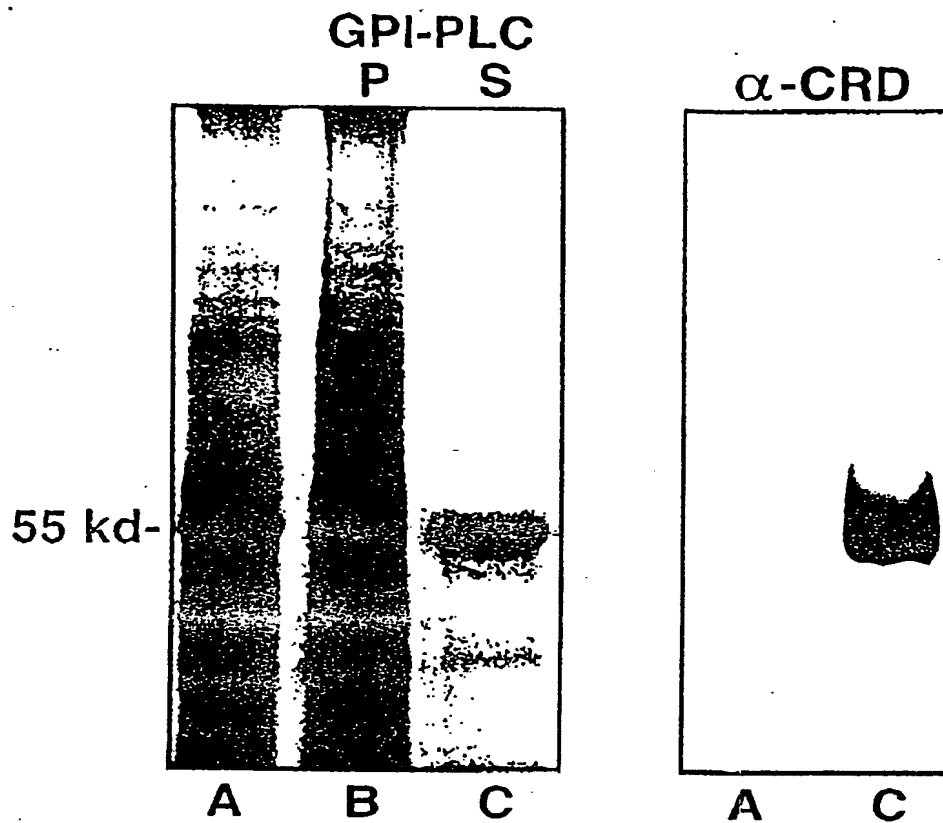


Fig. 11

SEB10 Primers for synthesis of G5 synthetic gene.

- NO. 70 3201:  
ATG GGA ATT CAA ATG AAG AAC AAC ATC CTG GTG ATC CTG ATC ATC TCT CTG TTC ATC AAC CAG ATC AAG  
TCT GCT AAC TGT CCT GTG GGA AOC GAG ACC AAC ACC GCT GGA CAG GTG
- 71 3202:  
CTC CAG GCA CGA AAG CAG CAG CGT TGT TGT AGT AGA AGT TCT TCT GAC AGT TCA CAC AGT TAG CAG GGG  
TTC CCA GGT CGT CCA OCT GTC CAG CGG TGT TGG TC
- 72 3203:  
CGC TGC TGC TTT CGT GCC TGG AGC TTC TAC CTG TAC CCC TTG TCC TCA GAA GAA GGA CGC TGG AGC TCA  
GCC TAA CCC TCC TGC TAC CGC TAA CCT GGT G
- 73 3204:  
GAT GAT AGC AGC GTA GTC GGT AGC TCC TCC AGC GAT AGC GGT TCC AGC AGG ACA CTT CAC GTT ACA CTG  
GGT CAC CAG GTT AGC GGT AGC AGG AG
- 74 3205:  
GCT ACC GAC TAC GCT GCT ATC ATC ACC GAG TGT GTG AAC TGT CGC ATC AAC TTC TAC AAC GAG AAC GCT  
CCT AAC TTC AAC GCT GGA GCT TCT ACC TGT ACC GCT TGT OCT GTG AAC CGC GTG GGA GGA GCT CTG ACC
- 75 3206:  
GGT GAA AGA GCG CAC GTA GTC GGT GGT CAC TCC GTC GTC CAG AGC GGT TCC GGT AGG ACA AGC CAC GTT  
ACA CTG AGC CAC GAT GGT AGC AGC GTT TCC AGC GGT CAG AGC TCC TCC CAC GCG
- 76 3207:  
GAC TAC GTG CGC TCT TTC ACC GAG TGT GTG AAG TGT CGC CTG AAC TTC TAC TAC AAC GGA AAC AAC GGA  
AAC ACC CCT TTC AAC CCT GGA AAG TCT CAG
- 77 3208:  
GTG ATG GTA GCG TCG TTT CCC AGG GTA GCC TGA GCC ACG TTA GCA GGC TTG ATA GCA GGA CAA GGG GTA  
CAC TGA GAC TTT CCA GGG TTG AAA GG
- 78 3209:  
GGG AAA CGA CGC TAC CAT CAC CGC TCA GTG TAA CGT GGC TTG TCC TGA CGG AAC CAT CTC TGC TGC TGG  
AGT GAA CAA CTG GGT GGC TCA GAA C
- 79 3210:  
CAG ACA GGT AGA GTT TCC AGG GTT GAA GTT AGG AGC GTT GTT GTT GTA GAA GTT AGG AGC ACA GTT GGT  
ACA CTC GGT GTT CTG AGC CAC CCA GTT GTT C
- 80 3211:  
CCC TGG AAA CTC TAC CTG TCT GCC TTG TCC TGC TAA CAA GGA CTA CGG AGC TGA GGC TAC CGC TGG AGG  
AGC TGC TAC CCT GGC TAA GC
- 81 3212:  
GGT CTG CAG GAT CAC GTA GTT GGT AGC TCC AGA AGC GAT AGC GGT TCC GTC AGG ACA AGC GAT GTT ACA  
CTG CTT AGC CAG GGT AGC AGC
- 82 3213:  
CAA CTA CGT GAT CCT GCA GAC CGA GTG TCT GAA CTG TGC TGC TAA CTT CTA CTT CGA CGG AAA CAA CTT  
CCA GGC TGG ATC TTC TCG CTG TAA GG
- 83 3214:  
GAG CGA TCA GGG TAG CGG TTC CTC CAG CGG TAG CCA CAG CTC CCT GCA CCT TGT TAG CAG GAC AAG CCT  
TAC AGC GAG AAG ATC CAG CCT GG
- 84 3215:  
GAA CCG CTA CCC TGA TCG CTC AGT GTG CTC TGG AGT GTC CTG CTG GAA CCG TGC TGA CCG ACG GAA CCA  
CCT CTA CCT ACA AGC AGG CTG CTT C
- 85 3216:  
GGT GTC GAT TCC AGC CAC CCA GTC GGT CTG CTT GGT GGT GTA GAA GTT AGC AGC ACA CTT CAC ACA CTC  
AGA AGC AGC CTG CTT GTA GGT AG
- 86 3217:  
GGG TGG CTG GAA TCG ACA CCT GTA CCT CTT GTA ACA AGA AGC TGA CCT CTG GAG CTG AGG CTA ACC TGC  
CTG AGT CTG CTA AGA AGA ACA TC
- 87 3218:  
GAG GGA TCC TTA TTA CAG CAG GTA GTA AGA GAT CAG CAG CAG AGA GAT AGA CAG GAA GTT AGC GAA GTC  
ACA CTG GAT GTT CTT CTT AGC AGA CT

Fig. 12



G5 proline mutant

Seq ID  
NO:53

10	↓ 20	30	40	50	60
ATGAAGAACA	ACATCCCGGT	GATCCTGATC	ATCTCTCTGT	TCATCAACCA	GATCAAGTCT
70	80	90	100	110	120
GCTAACTGTC	CTGTGGGAAC	CGAGACCAAC	ACCGCTGGAC	AGGTGGACGA	CCTGGGAACC
130	140	150	160	170	180
CCTGCTAACT	GTGTGAACTG	TCAGAAGAAC	TTCTACTACA	ACAACGCTGC	TGCTTTCTGT
190	200	210	220	230	240
CCTGGAGCTT	CTACCTGTAC	CCCTTGTCCT	CAGAAGAAGG	ACGCTGGAGC	TCAGCCTAAC
250	260	270	280	290	300
CCTCCTGCTA	CCGCTAACCT	GGTGACCCAG	TGTAACGTGA	AGTGTCTCTG	TGGAACCGCT
310	320	330	340	350	360
ATCGCTGGAG	GAGCTACCGA	CTACGCTGCT	ATCATCACCG	AGTGTGTGAA	CTGTCCGCATC
370	380	390	400	410	420
AACTTCTACA	ACGAGAACGC	TCCTAACTTC	AACGCTGGAG	CTTCTACCTG	TACCGCTTGT
430	440	450	460	470	480
CCTGTGAACC	GTGTGGGAGG	AGCTCTGACC	GCTGGAAACG	CTGCTACCAT	CGTGGCTCAG
490	500	510	520	530	540
TGTAACGTGG	CTTGTCCTAC	CGGAACCGCT	CTGGACGACG	GAGTGACCAC	CGACTACGTG
550	560	570	580	590	600
CGCTCTTTCA	CCGAGTGTGT	GAAGTGTCGC	CTGAACCTTCT	ACTACAACGG	AAACAACGGA
610	620	630	640	650	660
AACACCCCTT	TCAACCCCTG	AAAGTCTCAG	TGTACCCCTT	GTCCTGCTAT	CAAGCCTGCT
670	680	690	700	710	720
AACGTGGCTC	AGGCTACCCCT	GGGAAACGAC	GCTACCATCA	CCGCTCAGTG	TAACGTGGCT
730	740	750	760	770	780
TGTCCTGACG	GAACCATCTC	TGCTGCTGGA	GTGAACAACCT	GGGTGGCTCA	GAACACCGAG
790	800	810	820	830	840
TGTACCAACT	GTGCTCCTAA	CTTCTACAAC	AACAACGCTC	CTAACTTCAA	CCCTGGAAAC
850	860	870	880	890	900
TCTACCTGTC	TGCCTTGTC	TGCTAACAAAG	GACTACGGAG	CTGAGGCTAC	CGCTGGAGGA
910	920	930	940	950	960
GCTGCTACCC	TGGCTAAGCA	GTGTAACATC	GCTTGTCCTG	ACGGAACCGC	TATCGCTTCT
970	980	990	1000	1010	1020
GGAGCTACCA	ACTACGTGAT	CCTGCAGACC	GAGTGTCTGA	ACTGTGCTGC	TAACTTCTAC
1030	1040	1050	1060	1070	1080
TTCGACGGAA	ACAACCTCCA	GGCTGGATCT	TCTCGCTGTA	AGGCTTGTCC	TGCTAACAAAG
1090	1100	1110	1120	1130	1140
GTGCAGGGAG	CTGTGGCTAC	CGCTGGAGGA	ACCGCTACCC	TGATCGCTCA	GTGTGCTCTG
1150	1160	1170	1180	1190	1200
GAGTGTCTCTG	CTGGAACCGT	GCTGACCGAC	GGAACCACTT	CTACCTACAA	GCAGGCTGCT
1210	1220	1230	1240	1250	1260
TCTGAGTGTG	TGAAGTGTGC	TGCTAACTTC	TACACCACCA	AGCAGACCGA	CTGGGTGGCT

Fig. 13

proline mutant

1270	1280	1290	1300	1310	1320
GGAATCGACA	CCTGTACCTC	TTGTAACAAG	AAGCTGACCT	CTGGAGCTGA	GGCTAACCTG
1330	1340	1350	1360	1370	1380
CCTGAGTCTG	CTAAGAAGAA	CATCCAGTGT	GACTTCGCTA	ACTTCCTGTC	TATCTCTCTG
1390	1400	1410	1420	1430	1440
CTGCTGATCT	CTTACTACCT	GCTG.....	.....	.....	.....

Fig 13 (cont'd)

004020 23626460

G5 proline mutant protein

Seq ID  
No. 54

↓ 10 20 30 40 50 60

MKNNIPVILI ISLFINQIKS ANCPVGTETN TAGQVDDLGT PANCVNCQKN FYYNNAAAFV

70 80 90 100 110 120

PGASTCTPCP QKKDAGAQP N PPATANLVITQ CNVKCPAGTA IAGGATDYAA IITECVNCRI

130 140 150 160 170 180

NFYNNENAPNF NAGASTCTAC FVNRVGGALT AGNAATIVAQ CNVACPTGTA LDDGVTTDYV

190 200 210 220 230 240

RSFTECVKCR LNFYNGNNG NTFPNPGKSQ CTPCPAIPKA NVAQATLGND ATITAQCINVA

250 260 270 280 290 300

CPDGTISAAG VNNWVAQNTE CTNCAPNFYN NNAPNFNPGN STCLPCPANK DYGAETATAGG

310 320 330 340 350 360

AATLAKQCNI ACPDGTALAS GATNYVILQT ECLNCAANFY FDGNNFQAGS SRCKACPANK

370 380 390 400 410 420

VQGAVATAGG TATLIAQCAL ECPAGTVLTD GTTSTYKQAA SECVKCAANF YTTKQTDWVA

430 440 450 460 470 480

GIDTCTSCNK KLTSGAEANL PESAKKNIQC DFANFLSISL LLISYVLL.. ..

Fig. 14

004020" 29626460

004020" 29626460

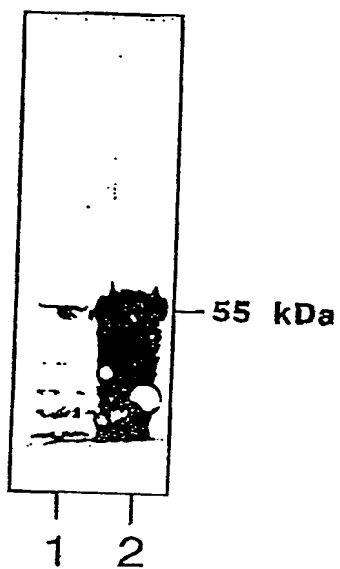


Fig. 15

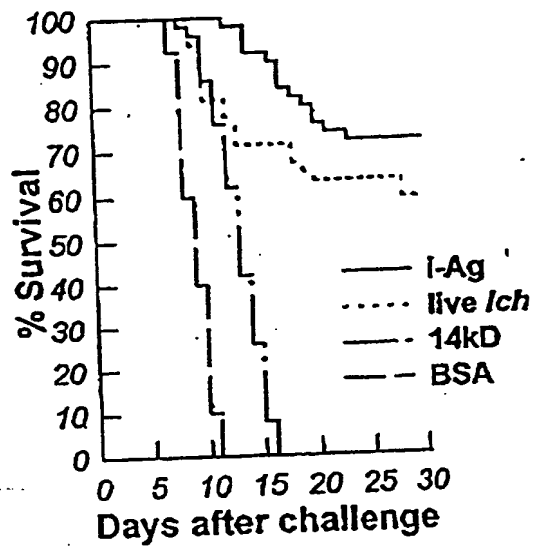


Fig. 14

004020" 29626460

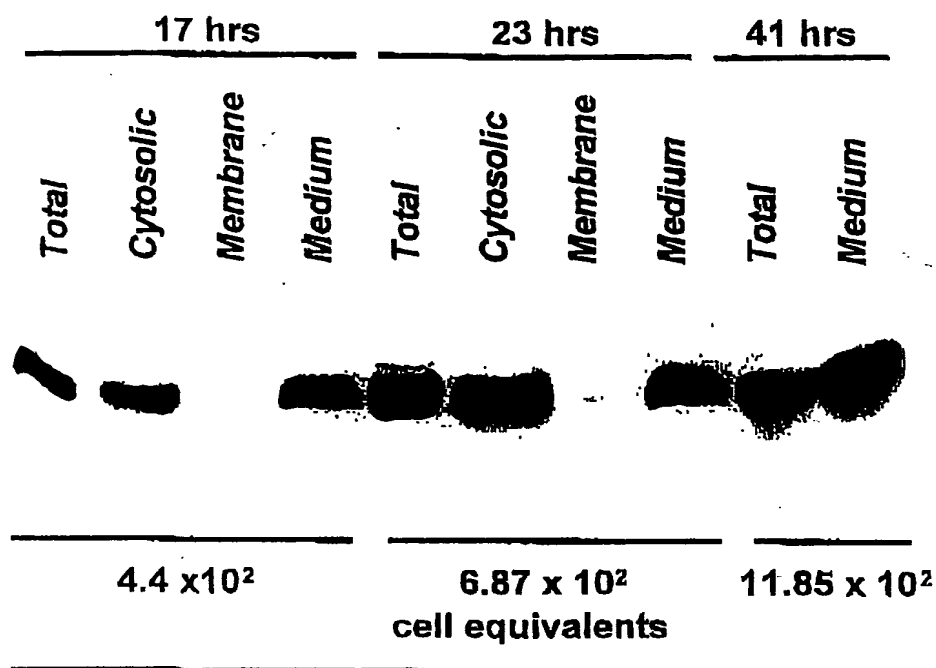


Fig. 17

004020" 29626460

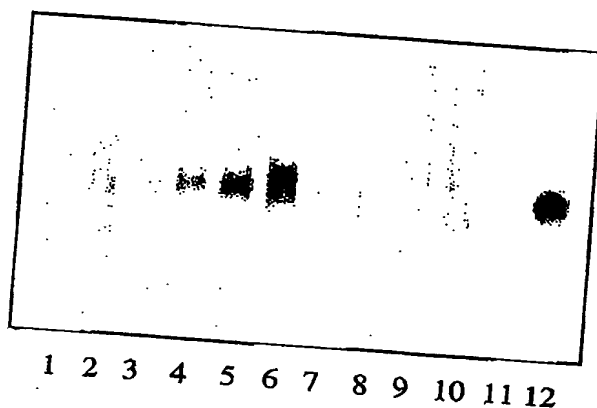


Fig. 18

004020-23626460

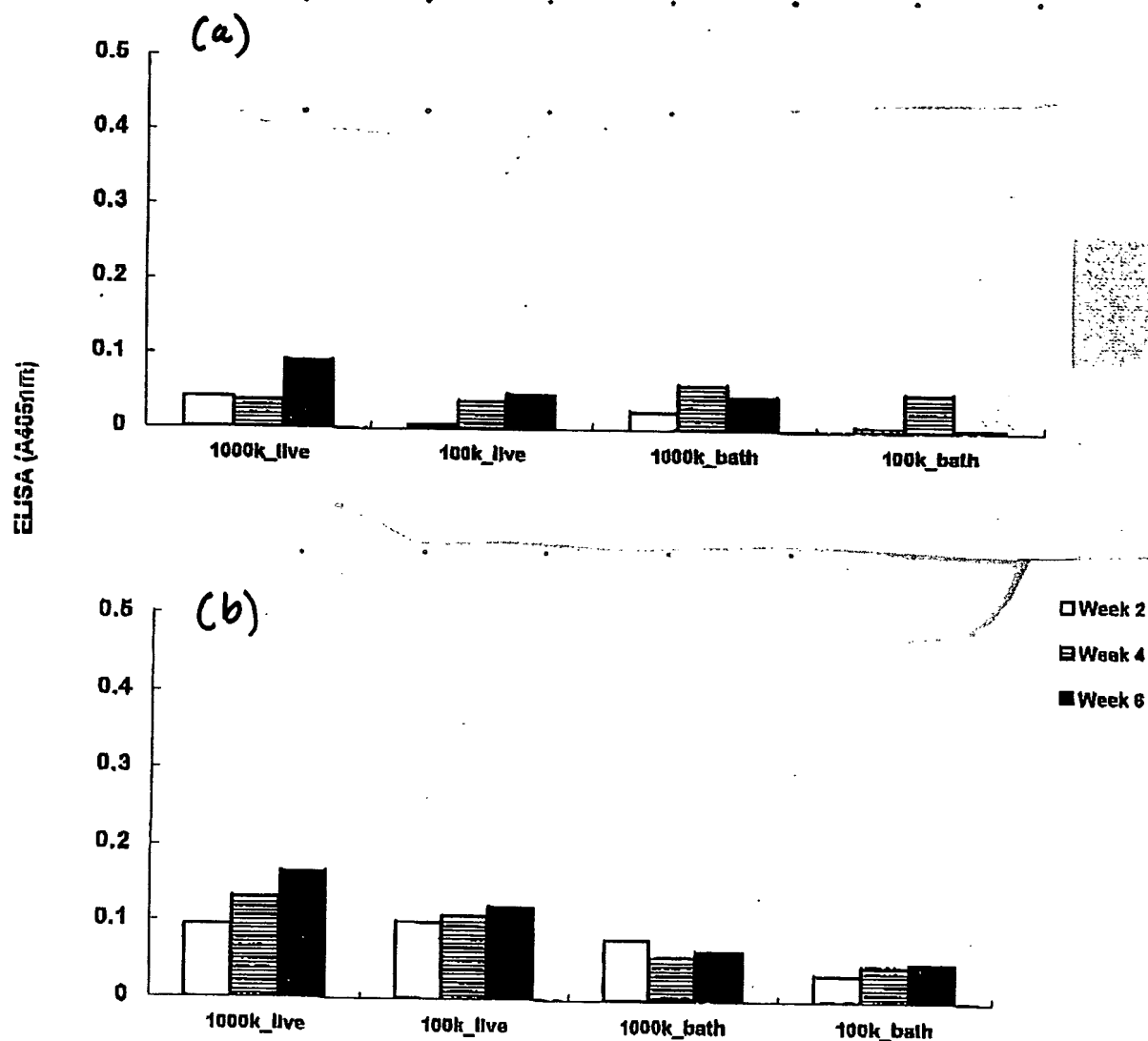
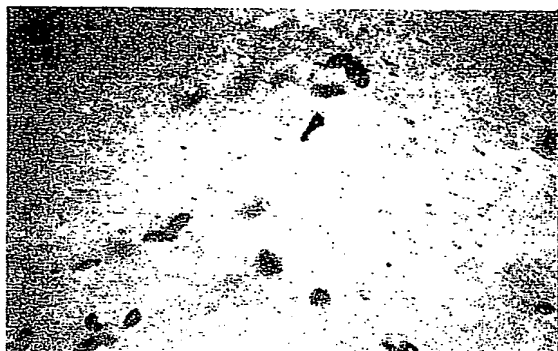


Fig. 19

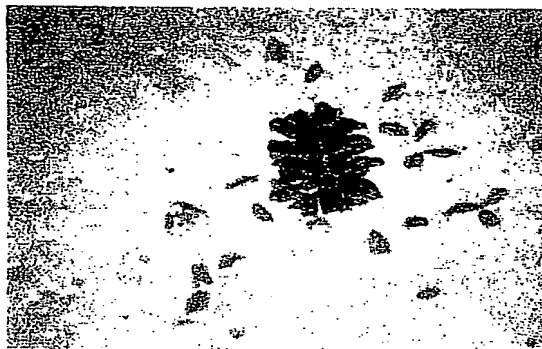


(a)



Serum: anti-live Tneo (1 : 20)  
(negative control)

(b)



Serum: anti-live TG1 (1 : 20)

Fig. 20

004020-29626460

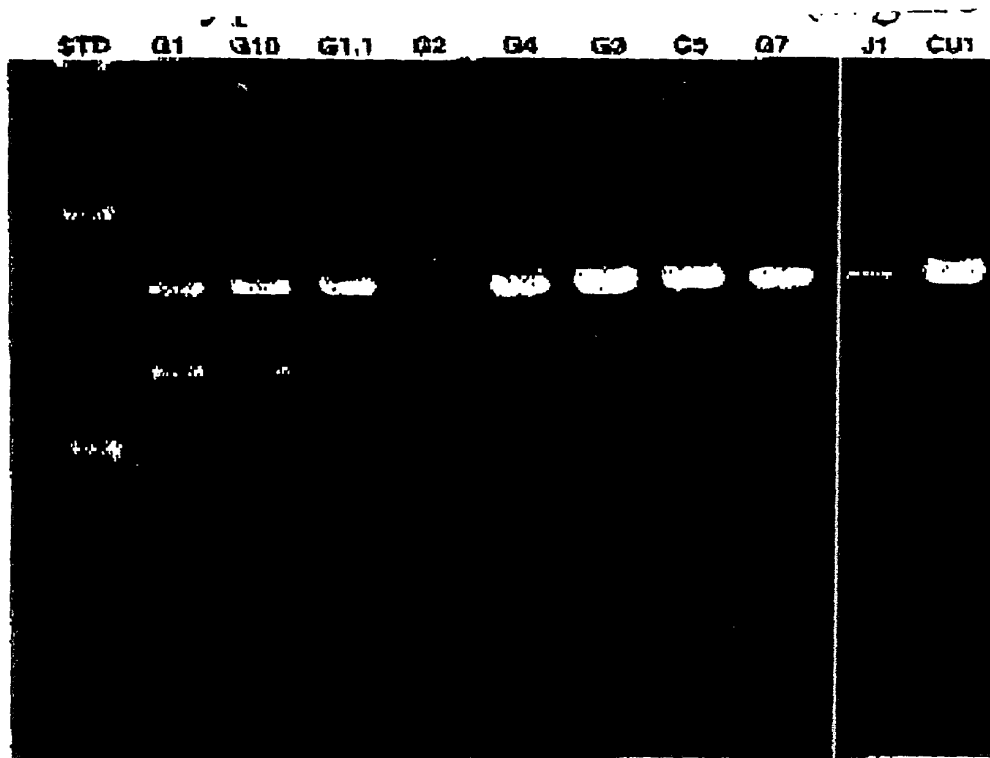


Fig. 24